1

ARI Research Note 90-120

AD-A227 063

Information Extraction From Visual Displays and the Event-Related Brain Potential

Chistopher D. Wickens, Carla Bosco, Amir Mane,
Arthur Kramer, Michael Coles, and Emanual Donchin

University of Illinois

for

Contracting Officer's Representative
Michael Drillings

Basic Research
Michael Kaplan, Director

September 1990

FIFTY YEARS OF SERVICE
1940 - 1990

United States Army
Research Institute for the Behavioral and Social Sciences

U.S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency Under the Jurisdiction of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Technical Director

JON W. BLADES
COL, IN
Commanding

Research accomplished under contract for the Department of the Army

University of Illinois

Technical review by

Nehama Babin

Accesion For	
NTIS CRA&I	☑
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution /	
Availability Codes	
Dist	Avail and / or Special
A-1	

DTIC COPY INSPECTED 1

NOTICES

DISTRIBUTION: This report has been cleared for release to the Defense Technical Information Center (DTIC) to comply with regulatory requirements. It has been given no primary distribution other than to DTIC and will be available only through DTIC or the National Technical Information Service (NTIS).

FINAL DISPOSITION: This report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The views, opinions, and findings in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other authorized documents.

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS —
2a. SECURITY CLASSIFICATION AUTHORITY —	3. DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE —	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) —	5. MONITORING ORGANIZATION REPORT NUMBER(S) ARI Research Note 90-120
6a. NAME OF PERFORMING ORGANIZATION University of Illinois Department of Psychology 6b. OFFICE SYMBOL (If applicable) —	7a. NAME OF MONITORING ORGANIZATION U.S. Army Research Institute for the Behavioral and Social Sciences
6c. ADDRESS (City, State, and ZIP Code) 603 East Daniel Street Champaign, IL 61820	7b. ADDRESS (City, State, and ZIP Code) Office of Basic Research (PERI-BR) 5001 Eisenhower Avenue Alexandria, VA 22333-5600
8a. NAME OF FUNDING / SPONSORING ORGANIZATION U.S. Army Research Institute 8b. OFFICE SYMBOL (If applicable) PERI-BR	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER MDA903-83-K-0255
8c. ADDRESS (City, State, and ZIP Code) 5001 Eisenhower Avenue Alexandria, VA 22333-5600	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. 61102B PROJECT NO. 74F TASK NO. NA WORK UNIT ACCESSION NO. NA

11. TITLE (Include Security Classification)

Information Extraction From Visual Displays and the Event-Related Brain Potential

12. PERSONAL AUTHOR(S) Wickens, Christopher; Bosco, Carla; Mane, Amir; Kramer, Arthur; Coles, Michael; and Donchin, Emanuel

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Interim	FROM 83/08 TO 86/09	1990, September	41

16. SUPPLEMENTARY NOTATION

Michael Drillings, contracting officer's representative.

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Displays, Event-related potentials, Decision heuristics, Information integration, Attention

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Three experiments, of increasing levels of complexity, are reported which examine two questions: (1) Will subjects extract more information from progressively more information stimuli in a problistic state-estimation task, as inferred from reaction time measures? (2) Will the amplitude of the P300 component of the event-related brain potential, reflect the amount of information extracted? The three experiments used different versions of a process monitoring task in which the process could be in one of two states, and information bearing on the expectancy of one state or the other was conveyed by discrete informative cues. Occasional probes signalled imperative responses to the expected or unexpected states. The data indicated that in the simplest version of the task with only two levels of information value (Experiment 1), both questions were answered affirmatively. In the more complex version with three levels (Experiment 2), one-half of the subjects performed in a manner consistent with the affirmative answers. In the most complex version (Experiment 3), in which information cues were spatially separated, and system states were autocorrelated, subjects did not extract information differentially from the different cues. The results are interpreted in terms of models of P300, decision effort and their applied implications for system design.

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL Michael Drillings	22b. TELEPHONE (Include Area Code) (202) 274-8722 22c. OFFICE SYMBOL PERI-BR

DD FORM 1473, JUN 86 — Previous editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

Information Extraction from Visual Displays and the Event Related Brain Potential

Christopher Wickens, Carla Bosco, Amir Mane, Arthur Kramer,
Michael Coles, and Emanuel Donchin

INTRODUCTION

Much of the information that humans typically process is of imperfect reliability and diagnosticity (Johnson, Cavenaugh, Spooner & Samet, 1973). Reliability describes the extent to which the value of an information cue that is displayed and perceived, reflects the true value of that cue. For example, if a thermometer is known to have an error of measurement, its reliability is reduced. Diagnosticity refers to the degree to which an information source, known to be reliable, can discriminate among several different states. If the thermometer is known to be reliable, it will be diagnostic in discriminating between the hyperthermia of heat stroke and the hypothermia of heat exhaustion, but will not be diagnostic in discriminating chicken pox from measles.

Together, in a multiplicative fashion, these two variables define information worth (I). "I" defines the degree to which the monitor of a system or state of affairs should update his or her belief on the basis of a cue or symptom (Schum, 1975). In the current design we describe I as the product /rxd/. Since reliability (r), and diagnosticity (d) may both be expressed as correlation coefficients, I, the absolute value of their product will vary between 0 and 1. When I = 0, the value of a cue contributes nothing to the understanding of system state. When I = 1, the single cue presents all the information there is to know about the system. Any further observations, taken at the same time, that are inconsistent must be in error.

There is an extensive body of research in human decision making that has focused on different normative or descriptive models for the revision of beliefs in the face of information of less than perfect worth (e.g., Edwards, Phillips, Hays, & Goodman, 1968; Edwards, 1977; Schum, 1975; Einhorn & Hogarth, 1985; Lopes, 1983; Slovic, Fischoff, & Lichtenstein, 1977), and there is a healthy range of opinions in this literature regarding the extent to which the human is, or is not, optimal. However, transcending the conclusions of all of these studies is one important principle. People should optimally extract more information from sources of greater worth than from those of lesser worth. Precisely how much more depends upon the nature of the models that are proposed as optimal.

This principle has been found to be violated in many laboratory inference tasks when subjects are required to integrate sources of varying information value, in updating their mental model of the state of a system. This tendency to treat all sources of information as if they were equally reliable has been described as the "as if" heuristic (e.g., Johnson, et al, 1973; Kahneman & Tversky, 1973; Wickens, 1984). To assert that people use the "as if" heuristic does not mean that they are unable to perceive these differences in cue properties when they are explicitly asked to do so, nor process them appropriately in the right kind of task. For example, an operator presented with one cue having an I of .5, and another with an I of .8, could easily treat these as "beta weights" and cross multiply the I values by the cue values occurring along each of the two channels to form an overall diagnostic probability. However, the findings of a substantial number of integration tasks indicate that subjects simply do not do so, in the context of multicue decision tasks. The reason for this departure from optimality is revealed by its label as a heuristic -- a rule of thumb that simplifies the cognitive operations required, and does work well most of the time. Thus, as shown in Table 1, most of the time, when the difference in information value of cues does not span a tremendous range, the answer provided by the "as if" heuristic will give a diagnostic outcome that is identical to that using a more optimal multiplicative strategy in which I's are optimally multiplied by cue values and then summed. One example of this convergence of the two strategies is shown in Table 1. However, this apparent optimality will break down as the number of hypotheses increase.

TABLE 1

		A	B	C	Net Evidence
1.	Cue Value*	-3.0	-2.0	+4.0	
2.	True I Weightings	.8	.6	.5	
3.	Product	-2.4	-1.2	+2.0	-1.6 (optimal)
4.	"As If" Weightings	.6	.6	.6	
5.	Product	-1.8	-1.2	+2.4	-0.6 ("As If")

*Sign indicates evidence for (+) or against (-) a hypothesis

Table 1 presents an example of three numerical cues bearing on the confirmation of one of two hypotheses (-) or (+). The table shows the effect of applying optimal weighting (row 2) and "As If" weighting (row 4) when integrating the values of three cues, A, B, and C, bearing upon the diagnosis of one of two hypotheses (-) and (+). Note that in this particular problem, as in the majority of others that might occur, the heuristic diagnoses the same state (-), as the optimal solution (although not by the same degree of confidence).

One major objective of the three experiments that we report here is to examine the "as if" heuristic. Can, or do people use the differing amounts of information provided by different cues, in a way to suggest that they employ an "optimum" weighting filter on the values of those cues?

A second major issue addressed by these studies is to determine if the P300 component of the event related brain potential (ERP) can index the amount of information extracted from sources of varying information value. This characteristic of the studies addresses points of both theoretical and applied importance. The theoretical interest spawns from an emerging model of the information processing routines underlying the generation of the P300 (Donchin, 1981). Donchin has proposed that the amplitude of the P300 reflects the amount of updating of working memory that is performed as a result of identifying the task-relevant properties of a stimulus. There are many characteristics of P300 that are consistent with such a model. For example, stimuli that are more surprising (and therefore force a greater reevaluation of the state of the world) produce larger P300s (Duncan-Johnson & Donchin, 1977). Stimuli that are not attended produce no P300s, and those that require progressively more resources to process generate larger P300s (Wickens, Kramer, Vanasse & Donchin, 1983; Kramer, Wickens & Donchin, 1985). However, evaluation of the collective set of results suggest that many of these results can be accounted for directly by variables of stimulus probability and attention, as well as by a context updating routine.

One purpose of the present studies therefore will be to present a series of stimuli that vary in their information value, but occur with equal probability. If stimulus probabilities are the primary variables responsible for variance in P300 amplitude, then such stimuli should elicit ERPs of equivalent amplitude. On the other hand, a context updating hypothesis directly predicts that cues of greater information value will require greater updating of the mental model of system state and therefore will produce larger P300s.

Two investigations provide evidence that is consistent with this viewpoint. In one study carried out by De Swart, Kok, and Das-Smaal (1981), subjects learned concepts by categorizing exemplars as either members or non-members, and received feedback regarding their response. The authors found that the amplitude of the late positive component, P300, was largest when the feedback stimuli conveyed task relevant information, prior to learning the concept defining rules. Once the concept rules were correctly hypothesized, during over-learning, the amplitude of the P300 measure declined. An investigation by Fabiani has shown that when subjects are presented a series of words, and engage in rote rehearsal of those words, those words which produce larger P300s upon presentation are more likely to be later recalled (Fabiani, Karis & Donchin, 1986). These results can be interpreted as if greater memory updating for these words produced both the larger P300 and the more permanent memory representation.

A second important theoretical point, addressed by the present investigation relates to the use of the P300 measure to help understand the "as if" heuristic itself. If subjects do fail to integrate differences in cue reliability, to what is this failure attributable? Is it simply that they have not processed the differences in the context of the inference task, or that they have processed them, but this processing is not reflected in the final behavioral measure of hypothesized state? This contrast may be likened to a similar dichotomy drawn between perceptual/cognitive limitations and response bias in accounting for subjects "conservatism" when updating Bayesian odds regarding the likelihood of different hypotheses (Edwards, 1968; DuCharme, 1970). In the current paradigms, the use of the "as if" heuristic will be inferred from behavioral (reaction time) measures. If such measures do reveal its employment then P300 amplitude will be examined to assess whether differences in cue information value may have been perceived, but simply not employed in revising the hypotheses which generated the response.

The potential applied value of the present research is reflected in terms of its relevance to display evaluation. As levels of automation increase in complex system design, human operators are more and more becoming monitors of extensive displays, with little overt activity (Wickens and Kessel, 1981). In evaluating such displays it would be very useful to have a passive measure of the amount of information extracted from display changes -- a measure that can be derived in the context of the ongoing task sequence, without the disruption caused by demanding an overt response. Oculometric measures of visual fixation (Harris & Spady, 1985; Wickens, 1986), provide only partial solutions because they are cumbersome to obtain, cannot discriminate attended from non-attended material within foveal vision, and cannot discriminate attention shifts between the auditory and visual modality. Evidence to date reveals that the P300 can nicely discriminate attended from non-attended display elements (Wickens, Heffley, Kramer, & Donchin, 1980; Kramer, Wickens & Donchin, 1983). The present research is intended to determine if more resolution can be obtained within the "attended" level of this dichotomy.

The present series of three experiments then are designed to assess whether P300 and behavioral measures of reaction time can index the amount of information extracted from a display, as the information value of a set of cues is varied. Three experiments present informative, preparatory cues followed occasionally by imperative information, for which a rapid response is required. Experiment 1 presents a very simple paradigm in which only two levels of information value are used. Informative stimuli have either some or no value in predicting which of two future states will be signalled. To the extent that subjects process the informative stimuli they will be able to prepare for a following "imperative" stimulus that is likely, and to respond rapidly to it when it occurs. To the extent that they do so, they will be financially rewarded. A correlation is that they should also respond particularly slow to an imperative stimulus for which they were <u>not</u> prepared (Posner, 1978).

Experiments 2 and 3 are more complex variations on the same theme, moving the experimental scenarios closer to the level of display complexity characterizing real world systems. In each of these studies there are again two possible states to which subjects must make appropriate and rapid responses at unpredictable occasions. However, in these experiments the diagnostic information, predictive of system state, may be presented at one of three levels: no information (50% predictive accuracy), 70% predictive information, and 90% predictive information. For example, after perceiving a cue that indicates with 90% certainty that the system is in state "A," the subject knows that <u>if</u> an imperative stimulus occurs and a response is suddenly required, nine times out of ten the stimulus will indeed reveal state A, and the response appropriate to this state should be given. Thus, it will be to the subjects' advantage to prepare to respond "A". The critical element that these experiments add to the first, is the presentation of <u>two</u> useful levels of information value. Can subjects diagnostically weight the more informative cue, more than the lesser cue? This discrimination is the point of potential breakdown manifest by the "as if" heuristic. While sharing this feature, Experiments 2 and 3 differ from each other in terms of their "real world" complexity. Experiment 2 explicitly presents the information value of each stimulus by a digit, and presents relatively unnatural random transition probabilities between each cue. Thus 90% state A cue, could be followed by a 90% state B cue with the same probability as a 70% State A cue. In reality, the first transition would be far less likely to occur than the second, since most systems have a high autocorrelation between consecutive readings of system state. Experiment 3 is embedded within the context of a more realistic process monitoring task. The information value of each cue is implicit in its location on the display, rather than explicit as a numerical value, and transition probabilities between cue values more closely approximate those of a true "sluggish" dynamic system.

EXPERIMENT 1

The paradigm used in Experiment 1 was adapted from Duncan-Johnson and Donchin (1982). In their study, subjects saw pairs of warning (S1) and imperative (S2) stimuli. The S1s varied, between blocks, in their predictive value concerning the S2s, and therefore allowed for more or less confidence in the response to S2. Duncan-Johnson and Donchin found that RT was faster to the S2s which were predicted by highly reliable S1s (confirming a strong expectancy) as compared to those following neutral S1s (which conveyed no information about the S2s). RT was slower to S2s which disconfirmed a strong expectancy. In addition, Duncan-Johnson and Donchin examined P300 amplitude and latency to the S2s, and found that larger P300s were elicited by S2s which disconfirmed a strong expectancy as opposed to a weak expectancy.

The present study sought to replicate and extend this work by examining P300 amplitude to the informative S1s, on the basis of the RT pattern elicited by the S2s. Our paradigm included mixed blocks of trials, in which the predictive value of the S1s varied from trial to trial. We reasoned that the P300 might provide a measure of the amount of information a subject was extracting from a stimulus, and predicted that the stimuli which carried more information would elicit a larger P300 than those which contained less information. Assuming that subjects extracted and used this differential information to their advantage, the speed of response to imperative stimuli was also expected to be faster following a stimulus with high information content than one that was less informative. In addition, we expected that an imperative stimulus that was consistent with the information presented by the informative stimulus would evoke a P300 with a smaller amplitude and a shorter reaction time than if the imperative stimulus was inconsistent.

Method: Experiment 1

Subjects

The subjects were three graduate students at the University of Illinois. All were right handed and had normal or corrected-to-normal vision. They were paid for their participation in the experiment.

Task

Subjects were seated 70 cm from a cathode ray tube (CRT) on which the stimulus letters H or S were briefly presented in the central location. The letters subtended a visual angle of three-quarters of a degree. Stimuli consisted of two types, informative and imperative. Informative stimuli were Hs or Ss with a dot at one of three positions to the right of the letter. The dots occurred either near the top of the letter (high), near the center of the letter (middle), or near the bottom of the letter (low). Imperative stimuli were Hs or Ss without a dot. These stimuli are shown in figure 1.1.

In different conditions, two kinds of sequences were presented. In the "fixed length" condition, subjects were presented with an informative stimulus followed by an imperative stimulus, and had to respond to an imperative H by pressing a button with one hand or an imperative S with the other. A high dot informative stimulus indicated with an 80 percent probability that the imperative stimulus would be the same letter (match) as the informative. Twenty percent of the time then the imperative stimulus was the other letter

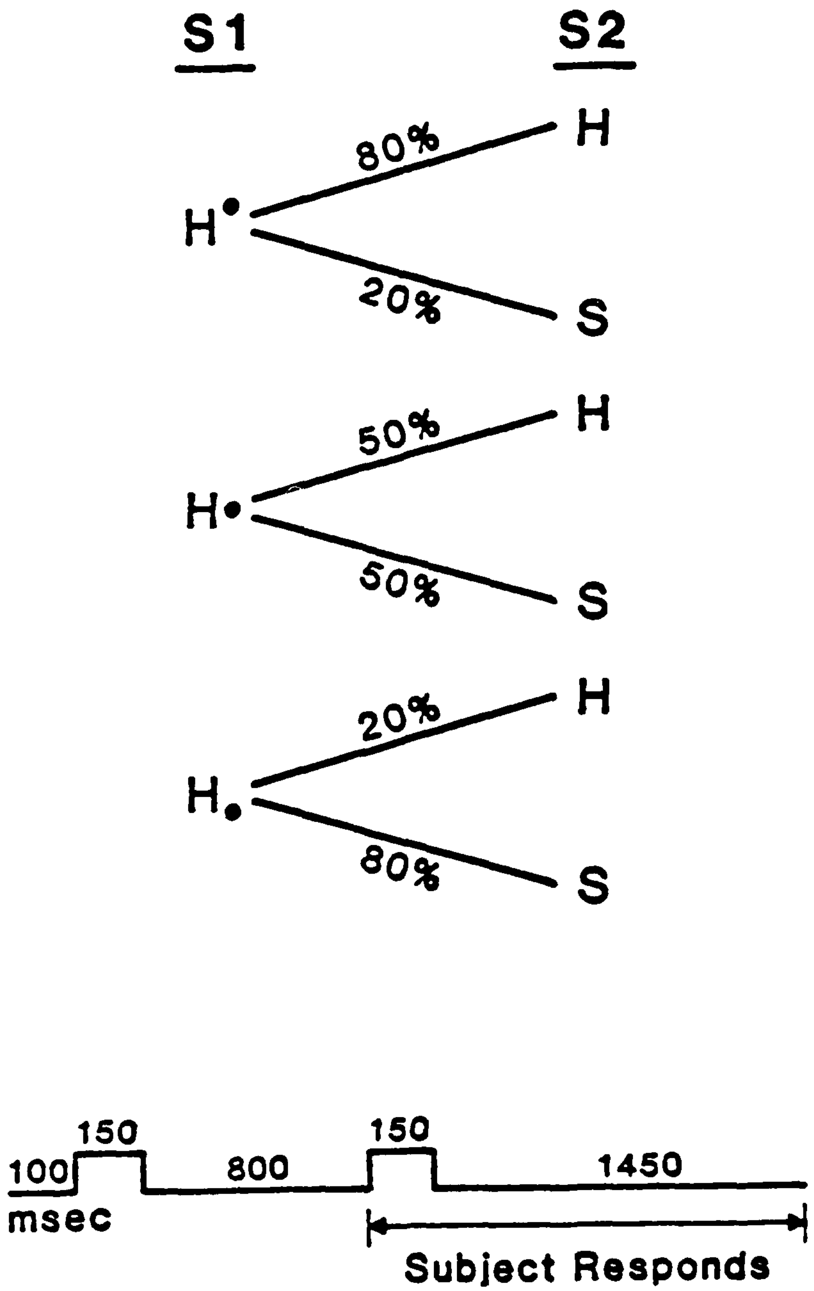

FIGURE 1.1. Stimulus sequence (S1-S2) and recording epochs. An identical set of S1 stimuli for the letter "S" could also be presented. Thus there were a total of six possible S1s, and twelve possible S1-S2 sequences.

(mismatch). When the dot appeared next to the middle of the informative stimulus, 50 percent of the following imperative stimuli were the same letter, and 50 percent were the other letter. That is, the "informative" stimulus in fact conveyed no predictive information. An informative stimulus with a low dot signalled a 20 percent likelihood that the imperative stimulus would be the same letter, while 80 percent of the following imperatives would be the other letter. Thus the information value of this stimulus, like the high dot stimulus was also .8, but the physical stimulus that it predicted was the opposite. The possible combinations of imperative and informative stimuli for the informative "H," and the timing of the sequences are displayed in Figure 1.1.

In the "variable length" condition, subjects were presented with sequences of either two or three stimuli (sequence length was randomly chosen). The last stimulus was always imperative, as in the fixed-length condition, while the first was always informative. However on a random half of the trials a second informative stimulus was presented. The identity of this second informative stimulus (H or S) was dictated by the dot probability of the first informative stimulus. The dot value of the second (80%, 50% or 20%) was randomly selected. Thus following S1, subjects had no knowledge of whether they would need to respond or simply would receive more information. Figure 1.2 illustrates the sequence. The purpose of this condition was to embed the informative value stimuli in the context of longer sequences, in a way that more closely approximated the conditions of Experiments 2 and 3.

It is important to note that there are two types of matches that can occur between the informative and imperative stimuli. As the term has been used above, match may refer to the physical identity between the two stimuli. Henceforth, this will be explicitly termed "physical match." The other meaning concerns the subjects' expectations of the imperative stimulus given the letter and dot level of the informative stimulus. This "expectancy match" may coincide with the physical match; when an H high dot S1 occurs, the presentation of an imperative H confirms both types of matches by validating the expectation as well as by being physically identical to the informative stimulus. An expectancy match-physical mismatch can occur when an H low dot S1 is followed by an imperative S--the subject <u>expects</u> the S2 to change, but there is a physical mismatch. Conversely, a physical match-expectancy mismatch will result from an H low dot informative paired with an H imperative.

Procedure

Each subject received detailed instructions as to the meaning of the dots, and was given practice with blocks (100 S1-S2 sequences) of each of the dot levels. The remainder of the first two sessions consisted of blocks in which the dot level varied within the block. On the third day, the subjects received blocks of the variable length sequences. Nine blocks were presented during each session. After each block, the subjects were informed of their error rate and speed of response. Speed instructions were emphasized. They were also told that each sequence was independent of the previous one. Subjects were free to request a break at any time, and if they had not exercised their option by the middle of a session, the experimenter imposed a "mandatory" break in order to reduce fatigue.

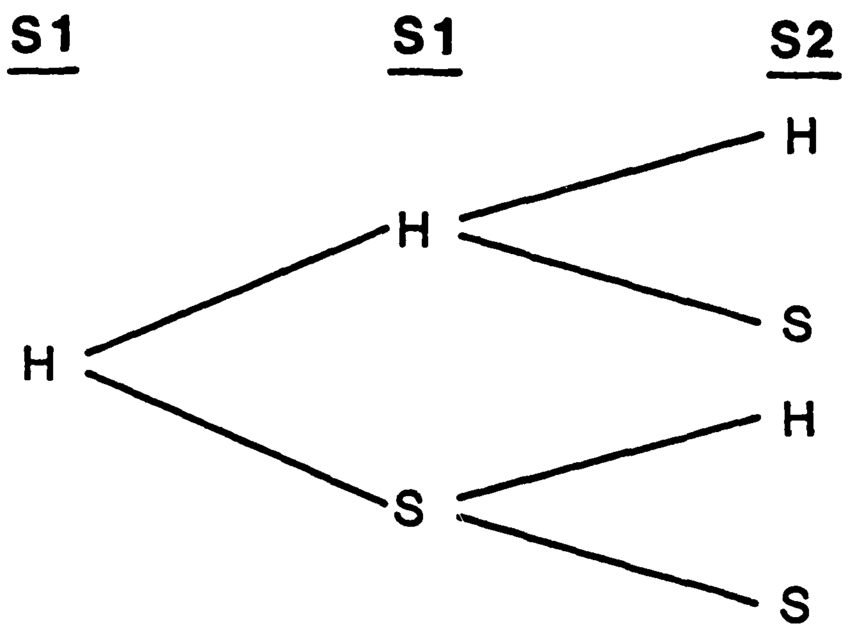

FIGURE 1.2. S1-S1-S2 sequences used in the variable length blocks.

ERP Recording and Analysis

ERPs were recorded from Fz, Cz, Pz, C3 and C4 electrode sites according to the International 10-20 system (Jasper, 1958) using Beckman Biopotential electrodes referenced to linked mastoids. Electrodes were also placed supra- and sub-orbitally to the right eye to record the electro-oculogram (EOG). ERPs and EOG were amplified by a Van Gogh model 50000 amplifier (ERPs: time constant = 10 seconds upper half amplitude of 35 Hz, 3db roll-off; EOG: time constant = 1 second, 15 Hz cutoff). Electrode impedances were kept below 10 Kohms. Both EOG and ERPs were digitized every 10 msec. The recording epoch was 1000 msec for the informative stimuli and 1600 msec for the imperative stimuli. Recording began 100 msec prior to the presentation of the informative stimuli, and 100 msec prior to the occurrence of the warning for the imperative stimuli. Single trials were adjusted off-line for significant EOG artifact by a regression-based correction procedure (Gratton, Coles & Donchin, 1983). All aspects of experimental control and data collection were controlled by a PDP-11/44 computer system interfaced with an Imlac graphics processor (Donchin and Heffley, 1975). Average waveforms and the single trial records were monitored using a GT44 display. Digitized single trial data and RT and accuracy of response on each imperative trial were stored on digital magnetic tape for later analysis.

Results: Experiment 1

Performance

The RT data of individual subjects were analyzed by computing means for six different categories. For each dot level (high, medium and low), RT was averaged according to whether the S2 matched the S1 or did not match (physical match) for the correct responses. Data were pooled across Hs and Ss. Separate averages were computed for the fixed and variable length blocks. The mean RTs for all subjects were submitted to a 3-way repeated measures analysis of variance (3 subjects x 3 dot levels x 2 match levels), and are plotted in Figure 1.3, which depicts the RT to imperatives as a function of the information value of the preceding cue. This information value is neutral (left, 50/50) or informative (right, 80/20 & 20/80). The lower limb of the figure represents the sequences in which expectancy was confirmed; the upper limb represents those when expectancy was violated. Solid lines connect the physical match pairs (HH or SS) and dashes lines connect physical mismatch pairs (HS, SH). The most prominent observation in Figure 1.3 is the RT advantage to expected stimuli over those that were either unpredicted (50/50) or counter-predicted (the expectancy mismatch) ($F=8.67$, $p<.05$). There is also a dot level x match interaction ($F=22.15$, $p<.01$) seen in Figure 1.3. This interaction, manifest in the crossing of the two lines of the upper limb, confirm that subjects were indeed processing the informative indicator within the stimulus. For example when a low dot informative S was, in fact, followed by an imperative S, a particularly strong penalty to RT was imposed. This is the physical and expectancy mismatch point in the far upper right of the figure.

Figure 1.4 presents the same pattern of results for the variable length sequences. There is a main effect for expectancy level ($F=57.05$, $p<.01$), further supporting the hypothesis that subjects are able to respond more quickly when presented with more information. The level x physical match

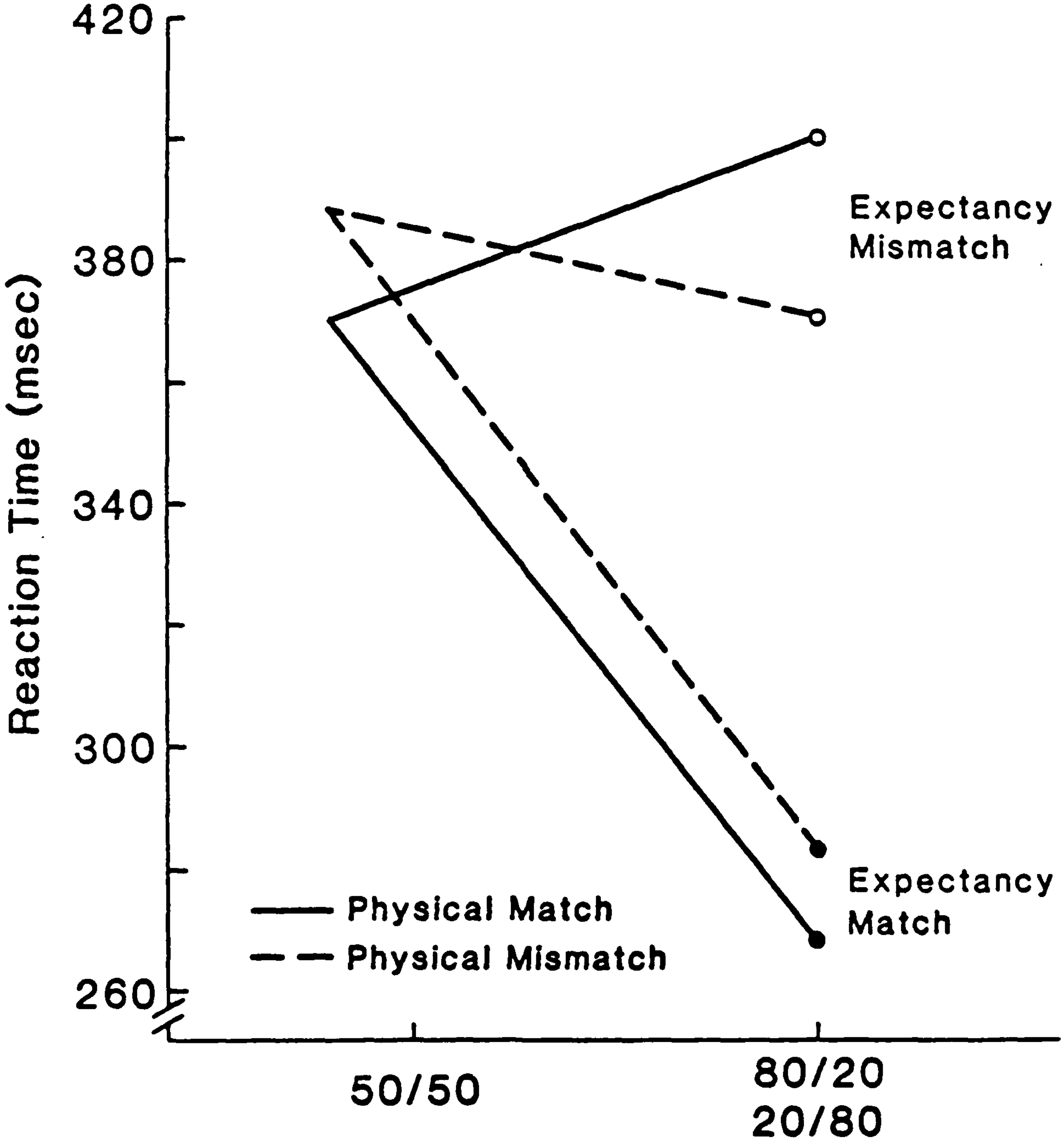

FIGURE 1.3. RT results for S1-S2 sequences showing main effect of dot level and dot level x match interaction

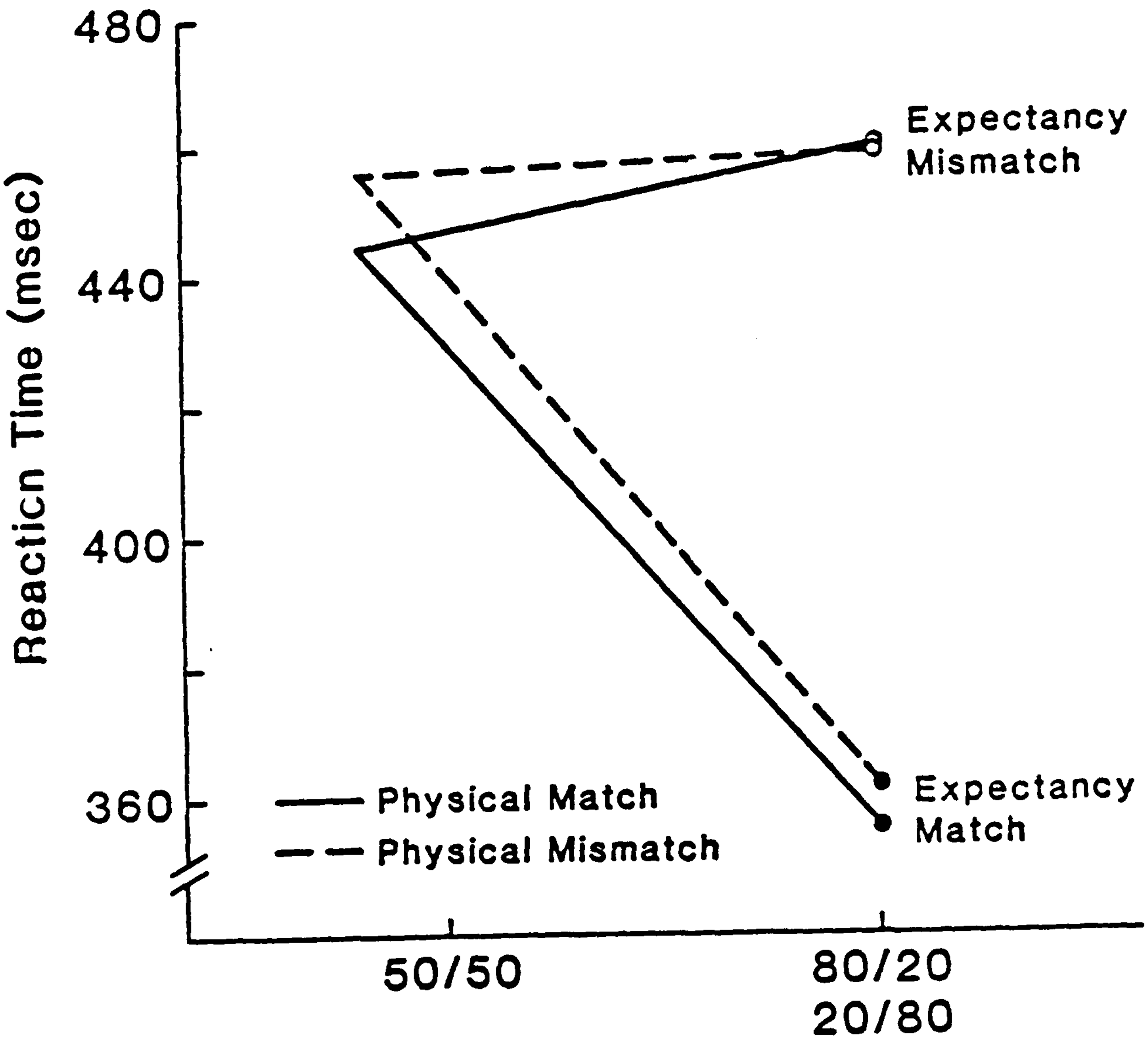

FIGURE 1.4. RT results for S1-S1-S2 (variable length) sequences

interaction is also significant ($F=21.62$, $p<.05$). Thus, in summary, the RT data verify that subjects did extract the relevant information from S1. Analysis of the ERP data assess how this extraction was manifest by brain processing.

Event Related Potentials

The ERP data from each subject were averaged separately for informative and imperative stimuli. Because of the small sample size, these data were not submitted to statistical analysis. However subsequent data analyzed while this report was in final preparation confirm the statistical significance of the effects described here.

For the informative stimuli, separate averages were computed for each electrode and for each dot level. The data from the Pz electrode are displayed in Figure 1.5. The data are plotted at Pz, since the P300 has been shown to be maximally positive at that site. The P300 is the large positive (down going) deflection of the waveform that begins about 300 msec after the stimulus (the vertical dashed line). It can be seen in the figure that the P300s elicited by the informative stimuli (high and low dots) are larger than those from the uninformative stimuli (middle dots depicted by he dotted line). Thus, visual inspection of the data supports the prediction that larger P300s are produced following more informative stimuli. This is true for both the fixed length sequences (one informative stimulus) and the variable length (one or two informative stimuli) sequences.

Figure 1.5 also reveals an apparent difference in latency between the 80/20 and 20/80 conditions. The longer latency was expected for the 20/80 condition due to the transformation required.

For the imperative stimuli, separate average ERPs were computed for each electrode site (Fz, Cz, Pz, C3, C4) for each dot level, and for match/mismatch. Figure 1.6(a) shows the waveforms from matching and non-matching imperative stimuli in the 80/20 condition at Pz. These data are averaged over informative stimuli from both the shorter and longer sequence. The P300 from the physical matches are smaller than those elicited by the physical mismatches. These results are consistent with the earlier findings that the amplitude of the P300 increases as expectancies are violated (Duncan-Johnson & Donchin, 1977, 1982). Figure 1.6(b) displays the waveforms obtained in the 20/80 condition. In this condition physical match and expectancy match are dissociated. That is, the 80% (expected) stimulus is a physical mismatch from the informative stimulus. The P300 elicited by the expected event which is also a physical mismatch is smaller than that elicited by the unexpected event which is also a physical match. These results are consistent with the earlier findings that the amplitude of the P300 increases as expectancy decreases (Duncan-Johnson & Donchin, 1977, 1982). As seen in Figure 1.6(c), the P300s to the matching and mismatching stimuli in the 50/50 condition are very similar to each other in amplitude.

Discussion: Experiment 1

The primary motivation for this study centered on the assertion that the P300 would indicate the amount of information that subjects extract from the informative stimulus, and that responses to imperative stimuli would be faster

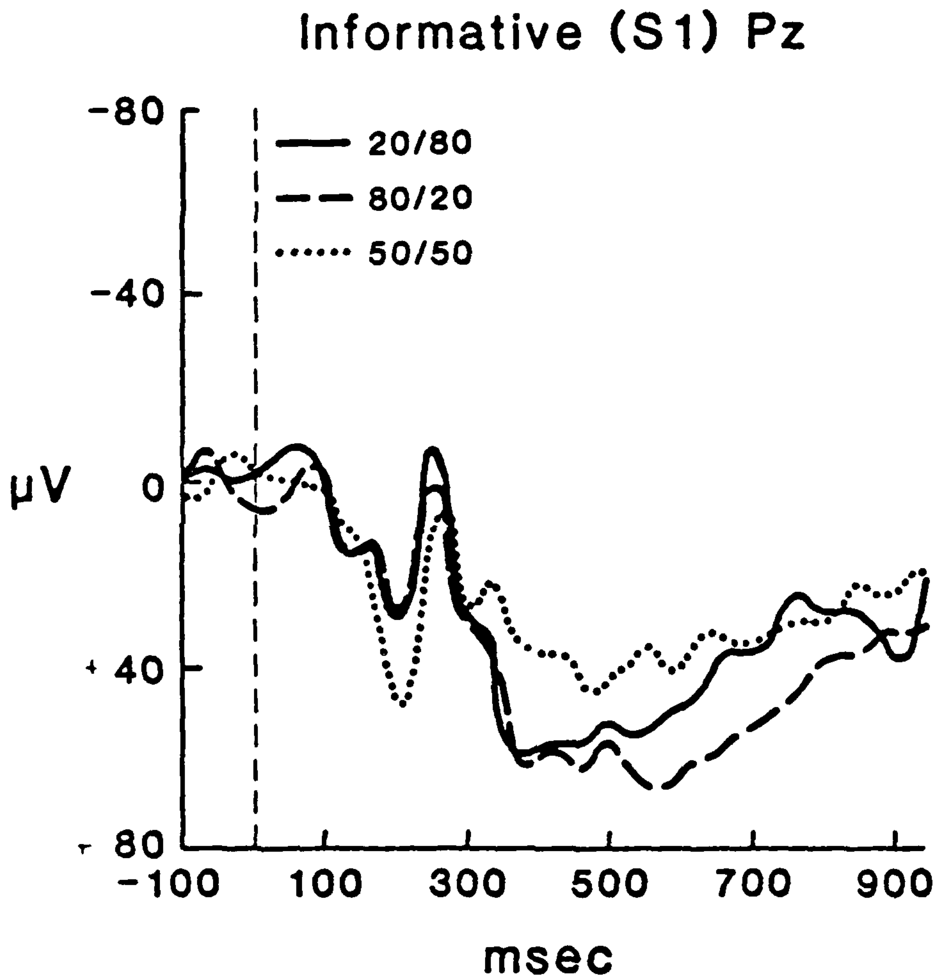

FIGURE 1.5. Grand Average 50/50, 80/20, 20/80 P300s informative stimuli

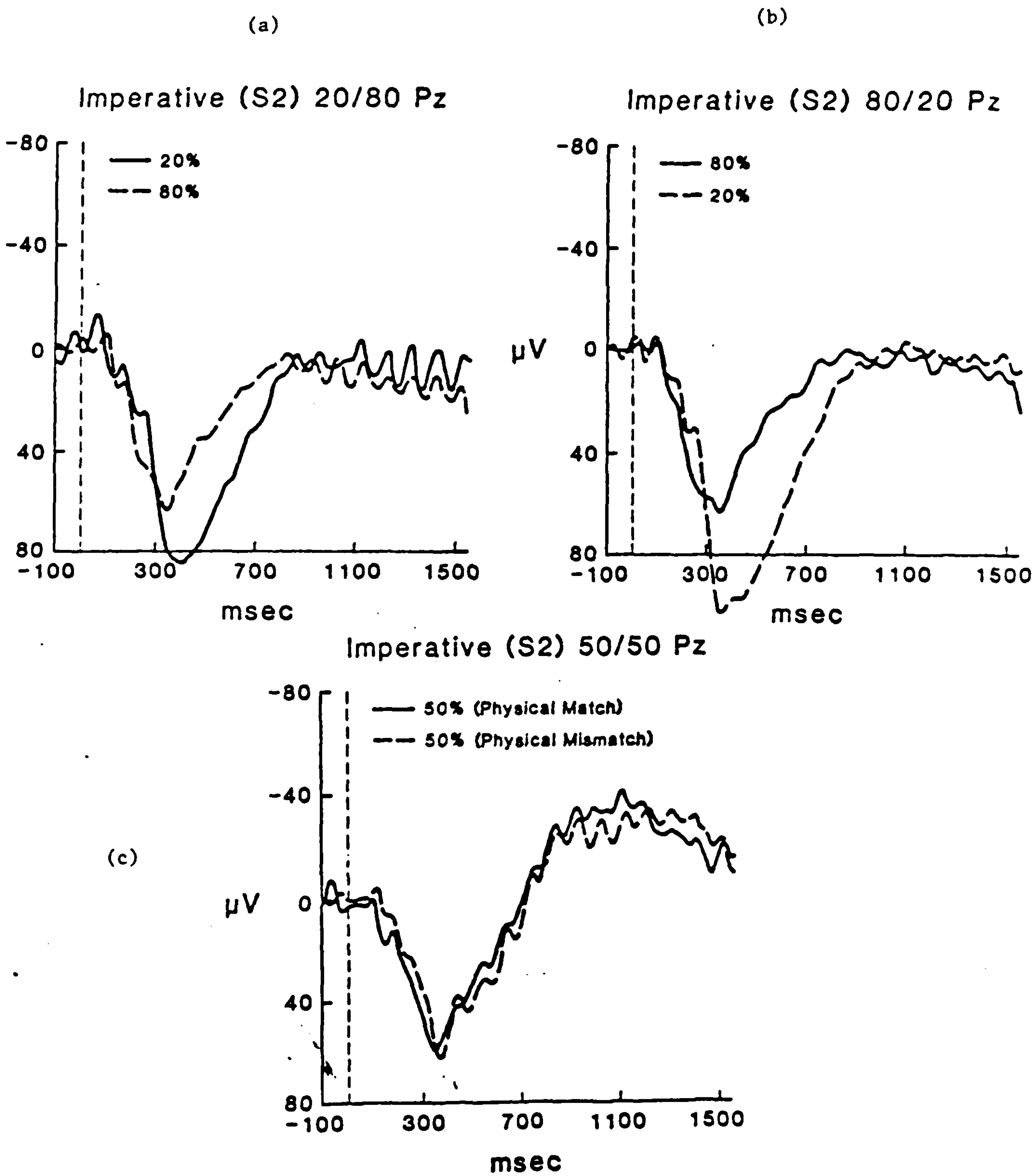

FIGURE 1.6. Average ERPs elicited by imperative stimuli that were preceded by a high dot (a), low dot (b), or medium dot (c) informative stimulus. The two lines within panels (a) and (b) indicate whether the imperative stimulus was expected (80%) or unexpected (20%), given the information in S1.

if those stimuli were predicted by highly informative cues. Both of these predictions appear to be supported by the data. RT was significantly faster following a highly informative cue than a less informative one. In addition, RT was faster when the imperative stimulus confirmed the expectancy given by the informative stimulus than when the expectancy was disconfirmed. The performance results show that subjects are indeed able to use the information provided to them.

The ERP results also confirm the predictions. The P300s from the more informative stimuli are larger than those from the less informative stimuli, although statistical tests, as mentioned above, were not employed because of the low power. A subsequent experiment, however, has replicated these results with statistical analysis.

EXPERIMENT 2

Experiment 2 was formally structured like experiment 1, but differed in the following important respects: (1) the sequences of informative stimuli were considerably longer, with greater uncertainty of the occurrence of an imperative stimulus; (2) three levels of information value corresponding to 0(.50-.50), .4(.70-.30) and .8(.90-.10) were employed; (3) a different display was used.

Methods: Experiment 2

Procedure and Design

Ten right handed subjects whose vision was normal, or corrected to normal, participated in the experiment. The subjects viewed a sequence of stimuli of three kinds: informative, warning, and imperative. In each sequence, the informative stimulus, which consisted of a shape (square for one hypothesis or circle for the other), circumscribing a digit (0, 4 or 8). The shape of the informative stimulus predicted the shape of the imperative stimulus which might follow. The reliability of the prediction was 50%, 70% or 90% for the three digits 0, 4 and 8 respectively. These numbers convey the advantage of using the information in the stimulus. Over a sequence of 20 trials the subject could expect to improve his accuracy over a guess by 8 trials if he uses a cue that predicts correctly 90% of the time. The other two numbers are derived in the same way.

Each informative stimulus was an independent event; its probability was not influenced by the preceeding informative stimulus. On 75% of the trials, the informative stimulus was followed by another informative stimulus with a different prediction. However, on 25% of the trials, the subject was presented with a warning signal (the letter x) which was followed 400 msec later by the imperative stimulus. This stimulus sequence is shown in Figure 2.1. Upon the presentation of the imperative stimulus, a square or a circle, subjects had to press a button corresponding to the presented stimulus. Subjects were rewarded monetarily for fast and accurate responses. Subjects were requested to prepare after each informative stimulus but were actually required to respond only when the entire sequence of 3 stimuli was presented.

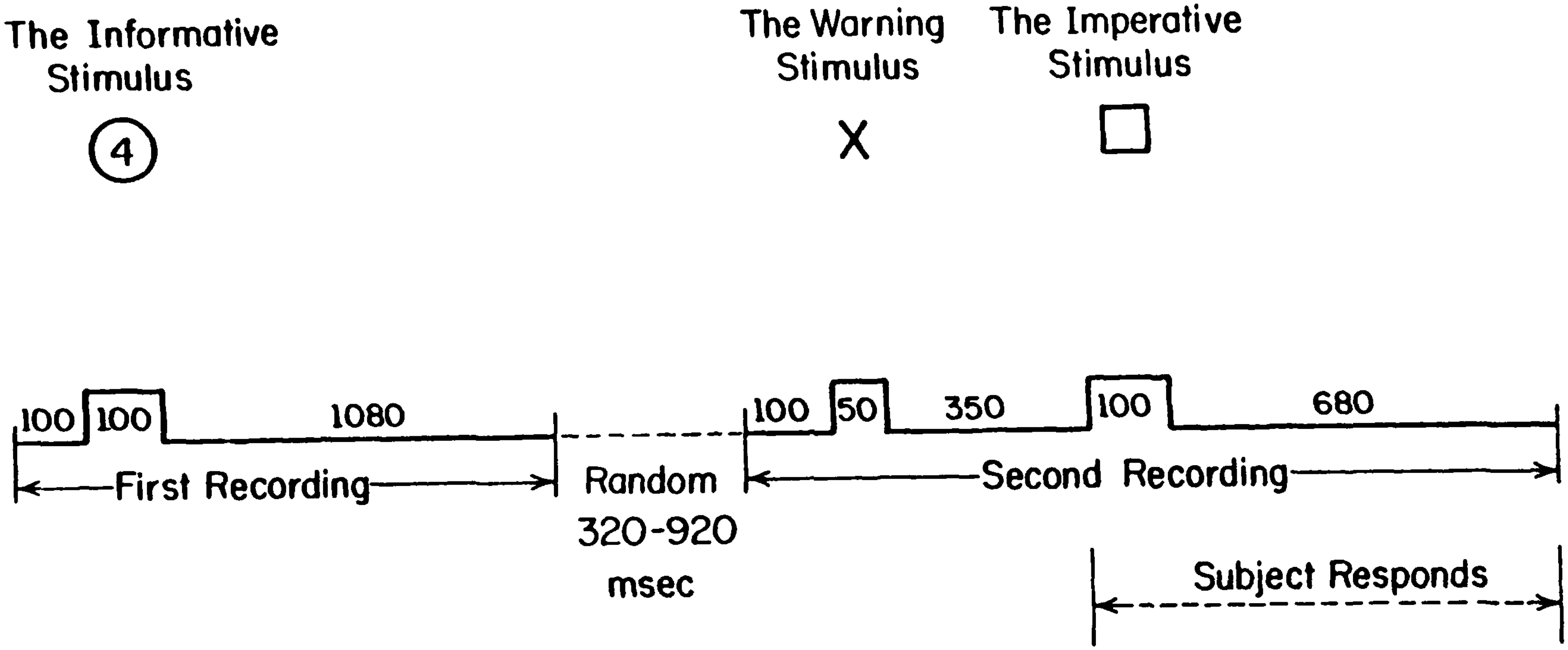

FIGURE 2.1. Sequence of events, showing informative stimulus, cue, and imperative stimulus. The circle state is presented in the informative stimulus in this example. Expectancy will be violated because the imperative is a square. This sequence is embedded within a largersequence of informative stimuli.

An important role of the instructions for the experiment was to teach subjects to use the information presented to them. Their training included the following: (1) Responding to the appearance of a square or circle by pressing the appropriate button (choice RT); (2) Using the prediction when there was 100% accuracy in prediction; (3) Using prediction when the reliability varied at levels of 50%-70%-90%. In these training conditions the only deviation from the experimental task was that the warning and imperative stimuli appeared every time; (4) Performing the experimental task as described earlier. During this training the imperative stimuli were presented on only 25% of the trials.

The last stage of training, identical to the experimental task, except that ERPs were not recorded, included 7 blocks of training consisting of 60 trials each. The experiment itself included 8 blocks of 60 trials. The total of 480 trials were presented, a quarter of which, 120 trials, included the full sequence of informative, warning, and imperative stimuli (20 of each of the six possible combinations).

ERP Recording and Analysis

The EEG was recorded from three midline and five lateral sites (Fz, Cz, Pz, C3 and C4 according to the 10-20 system) and referred to linked mastoids. Two ground electrodes were positioned on the left side of the forehead. Burden AG-AgCL electrodes affixed with collodion were used for scalp recording. Beckman Biopotential electrodes, affixed with adhesive collars, were placed laterally and supra-orbitally to the right eye to record EOG, and this type of electrode was also used for ground and mastoid recording. Electrode impedance did not exceed 10 Kohms/Cm. The EEG and EOG were amplified with Van Gogh model 50000 amplifiers (time constant 10 sec and upper half amplitude of 35 Hz). Both EEG and EOG were sampled for 1280 msec, beginning 100 msec prior to the informative stimulus onset and 100 msec prior to the warning stimulus onset. The data were digitized every 10 msec. ERPs were digitally filtered off-line (-3db at 8.8 Hz; 0db at 20 Hz) prior to statistical analysis.

Results: Experiment 2

Performance

As in experiment 1, an underlying assumption in the design of the task was that subjects would use the information presented in the informative stimuli. The expected data pattern would be to find faster responses to imperative stimuli that had been predicted with greater certainty (i.e., 90% vs. 70% and 70% vs. 50%). Correspondingly, slower reaction times should be shown to stimuli that were progressively less likely given the predictive information (i.e., an imperative stimulus that mismatched the shape of the previous informative stimulus). The reaction time data to the appearance of the imperative stimulus indicated that only five of the ten subjects treated the probabilistic cues differentially and showed this pattern. Figure 2.2 shows the RT for these subjects. In the case of an expectancy match (lower limb) it is clear that the higher the information value of the warning stimulus, the faster was subjects' response to the imperative stimulus. In the case of the expectancy mismatch (upper limb) the data are less consistent. This inconsistency results from two factors: (1) there are few data points,

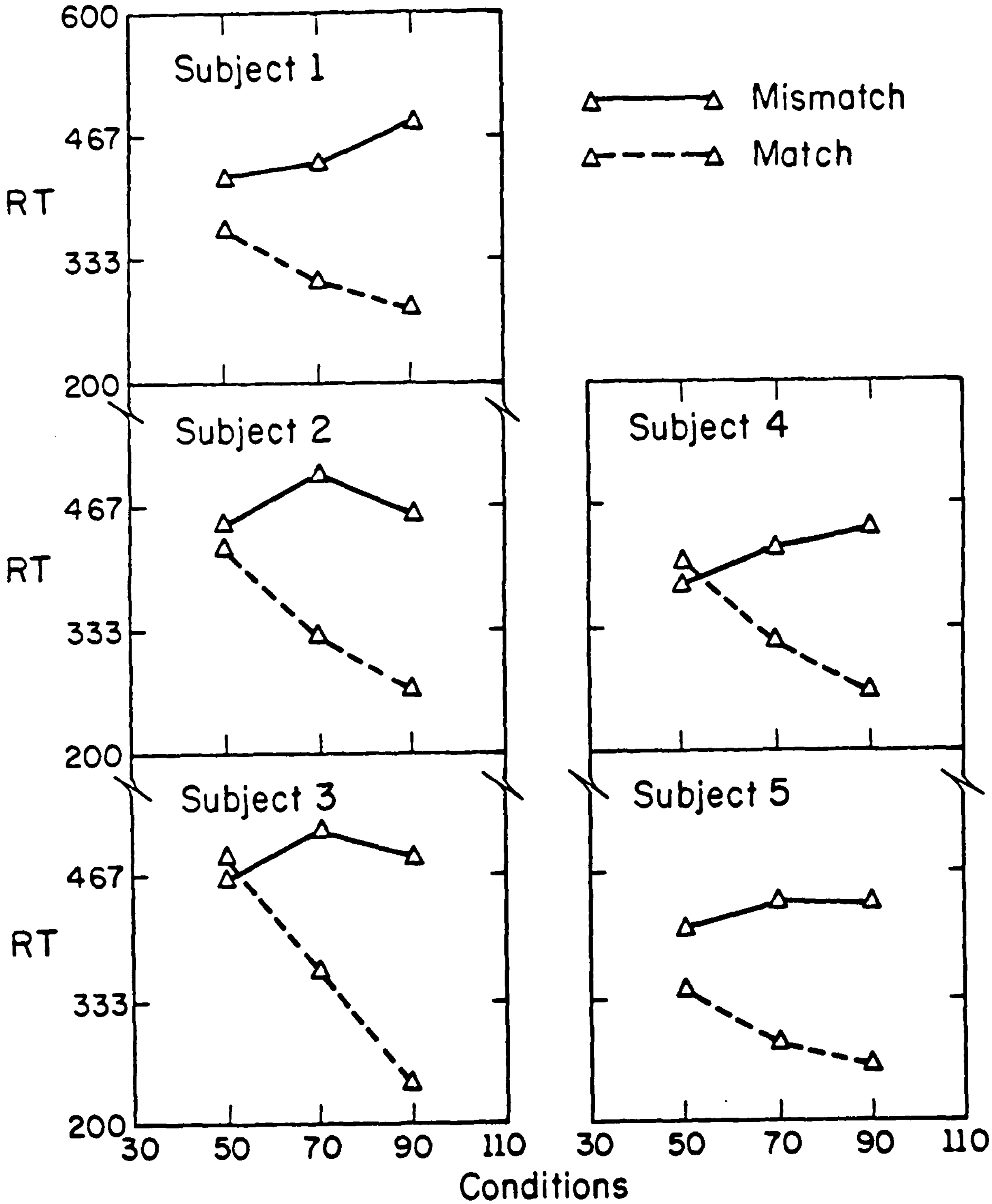

FIGURE 2.2. The average reaction times to imperative stimuli which either match or mismatch the informative stimulus. The lower limb - response to a match indicates the relative advantage of using the informative stimulus.

since the upper data point only occurred on 10% of the imperative trials following an "8" information value; and since "8" only occurred on 33% of the imperative trials, this data point only recorded for 3.3% of the total data base; (2) many of these trials were fast (and incorrect) guesses. The behavioral data for the other five subjects gave little indication that they were processing the informative stimulus, (no differential RT across information value). Hence the ERP data for these subjects were not examined.

Event Related Potentials

Figure 2.3 presents the ERPs elicited by the informative stimulus recorded at the Pz electrode and shows the large positive deflection of the waveform, following stimulus elicitation characteristic of the P300. Consistent with the hypothesis, P300 amplitude varied in a graded fashion with the information value of the cue. Principal component analysis of the recording indicated that the effect of information value on P300 amplitude is significant. However, this effect is due only to the greater amplitude in the 90% condition. The difference between the 70% and the 50% conditions was not significant.

Discussion: Experiment 2

The present results indicated that variance in P300 amplitude can be directly attributed to information value independently of probability. Performance data indicated that more predictive information was extracted from the 90% cue, and indeed, ERP amplitude was larger for these stimuli. These data support the conclusion that the P300 reflects the post-perceptual operations of memory update and information storage. However, certain aspects of the results did not yield a consistent pattern across all subjects and all levels of information.

Thus, on the one hand, only five subjects of the ten that were run in the experiment demonstrated the expected pattern of reaction time data. Therefore, it is necessary to ask why some of the subjects did not use the information. This failure is particularly salient given the elaborate training regime and the incorporation of financial incentives to use the information.

One explanation is that many subjects may have perceived a cost associated with extracting the information, which more than outweighed the small financial benefits realized by its use. An alternative possibility is that some portion of the subjects were simply incapable of employing the differential information value in a useful manner in this inference task, here manifesting use of the "as if" heuristic. Whether these subjects would adopt more optimal strategies if financial incentives were increased cannot, of course, be determined at this point. From the point of view of data analysis, however, we considered it appropriate to analyze the ERPs data of only those subjects who had demonstrated the appropriate performance response. It is important to emphasize that the categorization of subjects into the two groups on the behavioral criterion was carried out prior to examination of the ERPs.

With respect to the ERPs themselves, the results partially indicated their sensitivity to the degree of information extraction; significantly larger P300s were elicited by the stimuli of high information value, than by those of medium or no value. The interesting characteristic of this absence of effect between the 70% and 50% value cues, is that there were reaction time differences between stimuli following the two cues. Therefore, the differential information was apparently processed between 50% and 70% stimuli. Its processing was simply not reflected in P300. One hypothesis to explain this result is that subjects adopted a two-stage strategy in evaluating the diagnostic value of the information. At the first stage cues were categorized into two levels of being either highly informative (8), or less so (4 and 0). This difference was reflected in P300. Subsequently, the further discrimination between low and no information value was carried out, and was reflected in the faster RT to the more predicted stimulus. However, this evaluation was carried out after the ERP recording epoch had terminated, and hence was not manifest in the electrophysiological data. Nevertheless it is worth noting that there was a small difference in ERP amplitude between the 0 and 4 level stimuli. It is possible that an experiment with greater power may have revealed such a difference to be statistically reliable.

In summary, the results of Experiment 2 indicate that in a paradigm of greater complexity than that of Experiment 1, many subjects are in fact able to discriminate and use three levels of information value, when that value is explicitly presented to them in the form of numerical cues. P300 can also discriminate high from moderate information extraction, but does not appear to provide the same resolution of differences between all three levels of information value, as does reaction time. In Experiment 3, we extend the complexity of the paradigm still further toward the realism of the multi-element system monitoring task. This is accomplished by two formal changes: Information value is now implicit in the spatial location of a display, rather than explicitly indicated by a numerical value, and the hypothesized system state is now a gradually changing one, with an autocorrelation function that is greater than zero.

EXPERIMENT 3

Experiment 3 elaborates on the previous two by presenting subjects with three levels of diagnostic information within the context of a display monitoring task. The task was imbedded into the context of monitoring the water level in a nuclear power plant. Here the spatial location of the information representing different physical meters or information sources, indicated its predictive value. As in Experiments 1 and 2, subjects could respond more quickly if they attended to the informative warnings.

Method: Experiment 3

Subjects

Eleven university students (6 males and 5 females) served as paid volunteer subjects. All were right-handed and had normal or corrected-to-normal vision. Ages ranged from 19 to 22 years.

Task

Subjects viewed a cathode ray tube (CRT) screen on which four octagons were continually presented. The octagons were arranged in a triangle, with

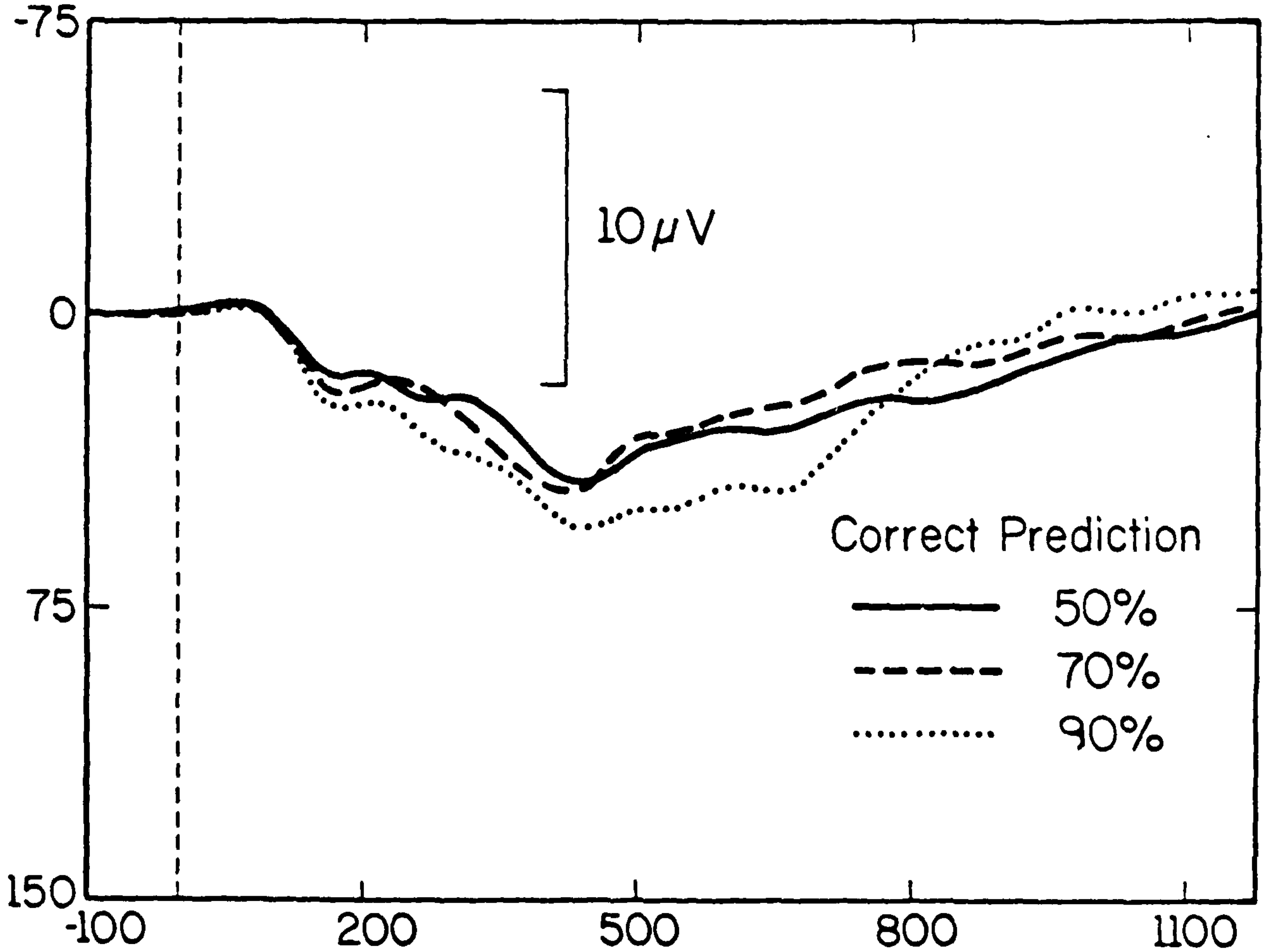

FIGURE 2.3. The average parietal waveform elicited by the informative stimulus. The three lines represent no (50%) medium (70%) and high (90%) informative value of the stimulus.

one at each vertex and one in the center. Arrows which pointed up or down could be presented inside the octagons, shown in Figure 3.1.

The three outer octagons served as the sources of information. The subjects were instructed to monitor the three gauges (the octagons) which indicated the water level in the cooling system of a nuclear power plant. Subjects were told that the gauges had different levels of reliability and were shown which octagon had the most (.9), the least (.5) and an intermediate level (.7) of reliability. The values 0, 4, and 8 have the same meaning as in Experiment 2, indicating the probability that the predicted state will agree with the actual state of .5, .7 and .9 respectively. The meaning of these values, and their association with display positions was carefully explained to the subjects. However the values was not displayed during the main experiment. The positions of each reliability gauge were counterbalanced across subjects.

Stimuli consisted of two types, informative and imperative. Informative stimuli were arrows presented in the peripheral octagons (the triangle vertices), and imperative stimuli were arrows which occurred in the center octagon. The direction of the arrows (up or down) indicated the water level in the hypothetical reactor.

A sequence of between three and seven informative stimuli were presented, followed by a flash of the perimeter of the inner octagon. As in Experiment 2, this flash served as a warning stimulus. Immediately after the warning, an imperative stimulus occurred. Subjects responded with a button press which indicated the state of the system (high or low water). Response hands were counterbalanced across subjects. A typical sequence of the three stimuli is shown in figure 3.2.

Each stimulus, whether informative, imperative or warning, was presented for 100 msec. The inter-stimulus interval was 1900 msec. The blocks were approximately eight minutes long. Within each block each information source was presented with equal frequency. The sequences of informational stimuli and imperative stimuli were generated by a transition probability table which made, for example, an up arrow in the 90 percent source more likely to be followed by an up arrow in the 70 percent source than a down arrow in the 90 percent source. This was done in order to achieve greater fidelity to real world systems, which usually change states gradually as opposed to oscillating randomly between all possible states.

Procedure

The subjects were seated in an unshielded room 70 cm in front of the CRT display screen, which was illustrated in Figure 3.1. Recording and control apparatus was located in a nearby room.

Each subject participated in two sessions on different days. The sessions consisted of eight blocks of 240 trials/block (trial = one stimulus, either informative or imperative). Each session lasted approximately 2 hours. After each block, the subjects were informed of their accuracy and speed of response. Speed instructions were emphasized. Subjects were also told that each sequence was independent of the previous one. After the third block on the first day, a bonus procedure was introduced in an effort to

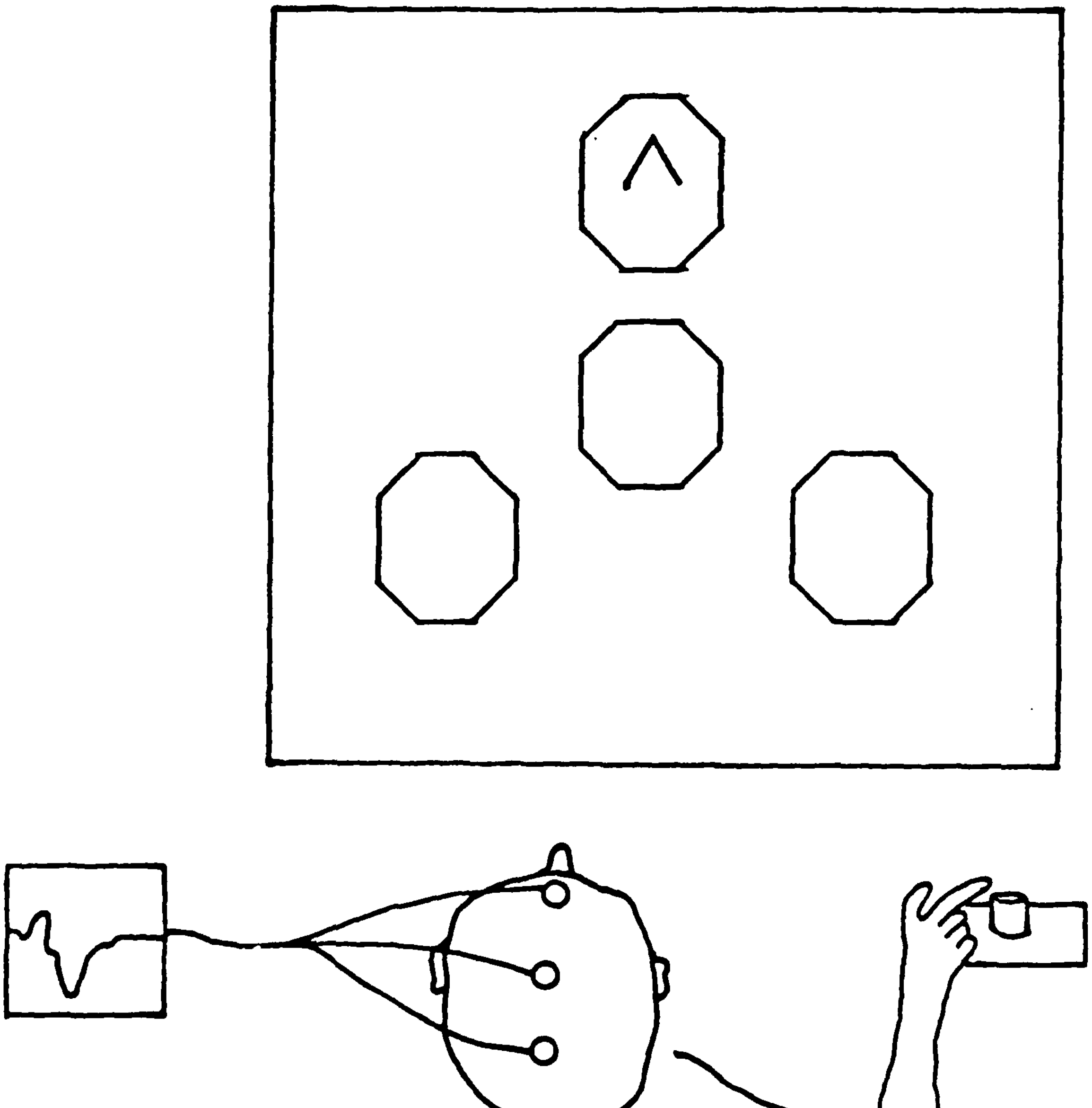

FIGURE 3.1. Subject at the display used in Experiment 3. An up arrow is present in one of the information sources, indicating that the water level in the system is high.

induce subjects to use the predictive information in a manner that was consistent with the experimental goals. Subjects were told that they could receive a bonus of 50 cents per block if their reaction times followed the correct pattern, characterized by subjects 1 and 4 in Figure 2.2, and if their mean RT either decreased or remained the same as the previous block. Subjects were free to request a break at any time, and if they had not exercised their option by the end of the fourth block the experimenter imposed a "mandatory" break in order to reduce fatigue.

ERP Recording and Analysis

ERPs were recorded from Fz, Cz, Pz, C3 and C4 according to the International 10-20 system (Jasper, 1958) using Beckman Biopotential electrodes referenced to linked mastoids. Electrodes were also placed supra- and sub-orbitally to the right eye to record the electro-oculogram (EOG). ERPs and EOG were amplified by a Van Gogh model 50000 amplifier (ERPs: time constant = 10 seconds upper half amplitude of 35 Hz, 3db roll-off; EOG: time constant = 1 second, 15 Hz cutoff). Electrode impedances were kept below 10 Kohms. Both EOG and ERPs were digitized every 10 msec. The recording epoch was 1280 msec for the informative stimuli and 2560 msec for the imperative stimuli. Recording began 100 msec prior to the presentation of the informative stimuli, and 100 msec prior to the occurrence of the warning for the imperative stimuli. A stimulus sequence is shown in Figure 3.2. Single trials were adjusted off-line for significant EOG artifact by a regression-based correction procedure (Gratton, Coles & Donchin, 1983). All aspects of experimental control and data collection were controlled by a PDP-11/44 computer system interfaced with an Imlac graphics processor (Donchin and Heffley, 1975). Average waveforms and the single trial records were monitored using a GT44 display. Digitized single trial data and RT speed and accuracy of response on each imperative trial were stored on digital magnetic tape for later analysis.

Results: Experiment 3

Performance

The RTs of individual subjects were analyzed by computing means for each information level (90%, 70%, 50%), for trials where the imperative stimulus matched the preceding informative stimulus, and for trials where the state of the imperative did not match that of the informative stimulus (mismatch). As in Experiment 2, it was expected that RT would decrease as a function of information level on the match trials and increase as a function of information level on the mismatch trials. This pattern of responding characterized the data of S's 1 & 4 in Experiment 2 (Figure 2.2.). This prediction may be compared with the pattern of actual data from Experiment 3 in Figure 3.3. This figure illustrates that the RTs to the mismatch trials are longer than to the match trials, and that there is a trend toward the expected interaction of information value with stimulus match.

A repeated-measures analysis of variance (ALICE statistical package, Grubin, Bauer & Walker, 1976) was conducted on three dependent variables separately. The dependent variables analyzed were mean RT, median RT and error rate as a function of information level. This led to three four-way ANOVAs (11 subjects x 2 sessions x 2 conditions [match vs. mismatch] x 3

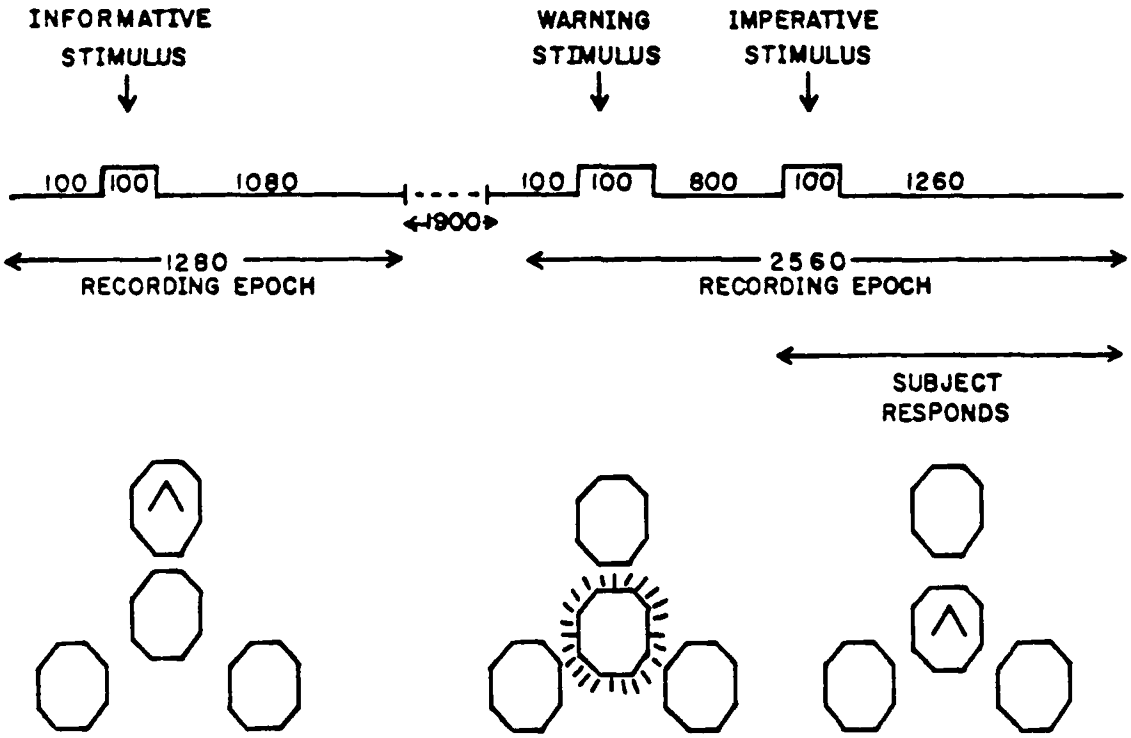

FIGURE 3.2. Typical stimulus sequence in Experiment 3 showing the relative timing of events. The particular sequence shown is one in which a level of expectancy represented by the top octagon (high water) is confirmed by the imperative stimulus.

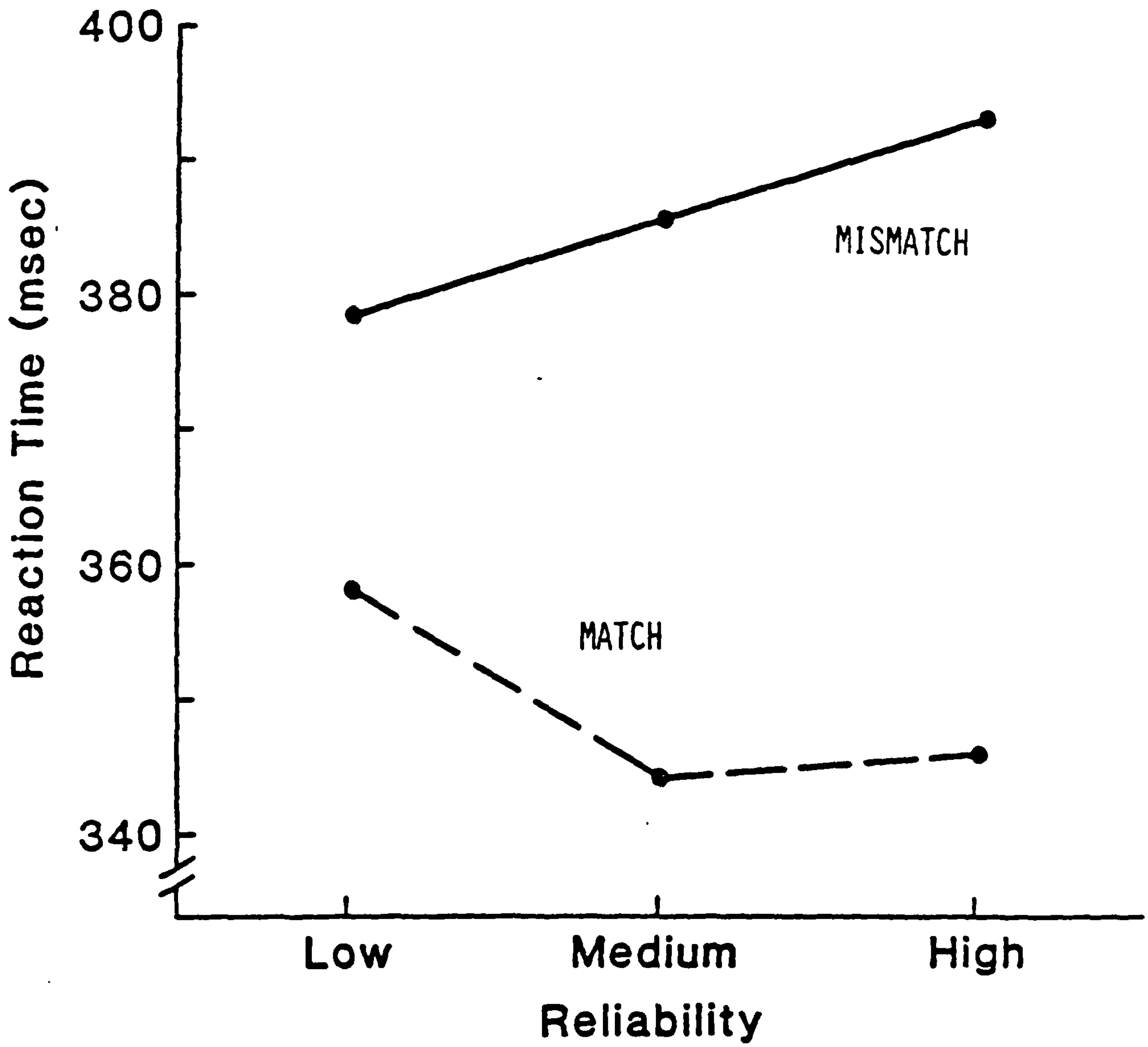

FIGURE 3.3. Graph of mean RTs in Experiment 3. Solid line connects responses to imperative mismatches, dashed line connects responses to imperative stimuli that were the same as the preceding informative stimuli.

information levels). The medians were analyzed in addition to the means because the RT distributions were skewed.

The ANOVA revealed that mismatch trials were significantly longer than match trials (means: $F(1,10)=14.80$, $p<.01$; medians: $F(1,10)=13.01$, $p<.01$), and that mismatches also produced significantly more errors ($F(1,10)=7.76$, $p<.05$). However in spite of the apparent trends in the data, the interaction between match and information value, depicted in Figure 3.3, which would have suggested greater use of the more informative stimuli, was not significant ($F(1,10)=2.32$, $p=.12$). Neither were significant linear trends revealed within the upper or lower "limbs" of the figure when examined separately. Reaction times were significantly shorter for session 2 than for session 1 (means: $F=5.70$, $p<.05$, medians: $F=13.57$, $p<.01$), indicating that subjects' performance was improving with practice.

Event-Related Potentials

ERP data from each subject were averaged separately for informative and imperative stimuli. For the informative stimuli, separate averages were computed for each electrode site (Fz, Cz, Pz, C3 and C4). A grand average across all subjects data was computed at Pz for the each level of information value to assess whether P300 was larger for the more informative stimuli. Figure 3.4 shows that there were essentially no differences in P300 amplitude as a function of information value. It appears that subjects were not differentially extracting the information available to them.

The next level of analysis focussed on single trials. Here, the question was whether, for a given information level, subjects extracted more information from stimuli that enabled them to respond more quickly, in a way that would produce larger P300's for these subjects. This question was assessed by dividing the RT distributions for each subject, and for each information level, into three categories, defined by the fastest, medium range, and slowest third of their RTs. The fastest third was labelled "fast" RT, the next third "medium" RT, and the last third "slow" RT. ERPs to informative stimuli were aggregated on the basis of the portion of the informative RT distribution with which they were associated. The results from this procedure are shown in Figure 3.5. It is apparent from the figure that no consistent patterns emerged.

The imperative stimuli were averaged separately for each electrode (Fz, Cz, Pz, C3 and C4) for each information level. Because the P300 component of the ERP has been shown to be maximally positive at Pz, the waveforms at this site were reaveraged by the information value of the preceding informative stimulus and according to whether the imperative stimulus was a match or a mismatch. A grand average across subjects is presented in Figure 3.6.

The P300 from the match conditions is smaller than that elicited by the mismatch stimuli. These results are consistent with the pattern of reaction time data in Figure 3.3, and with earlier findings that the amplitude of the P300 increases as stimulus expectancy decreases (Duncan-Johnson and Donchin, 1977, 1982). The absence of any amplitude differences across reliability level is also consistent with the reaction time data, indicating again that subjects failed to prepare differentially as a function of information value. Preparation was apparently carried out as a two-state process in which the

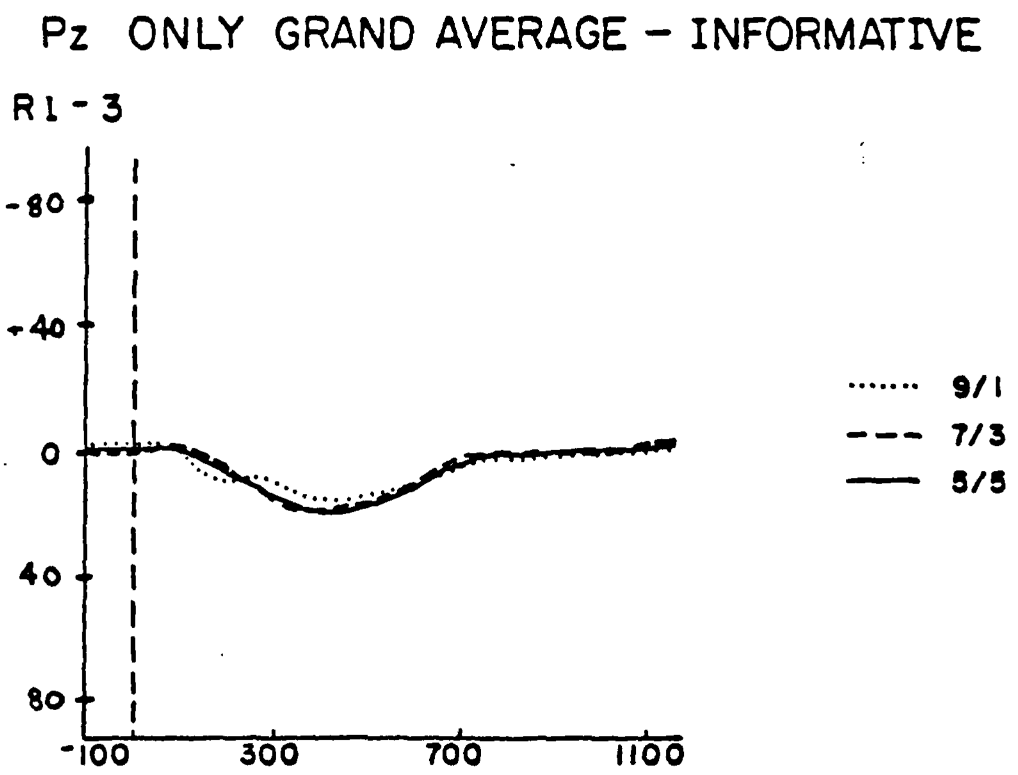

FIGURE 3.4. Average ERPs for informative stimuli at Pz for each information value. The dotted line represents ERPs from the 90% informative stimuli, 70% informative stimuli are plotted with a dashed line, and 50% informatives with the solid line.

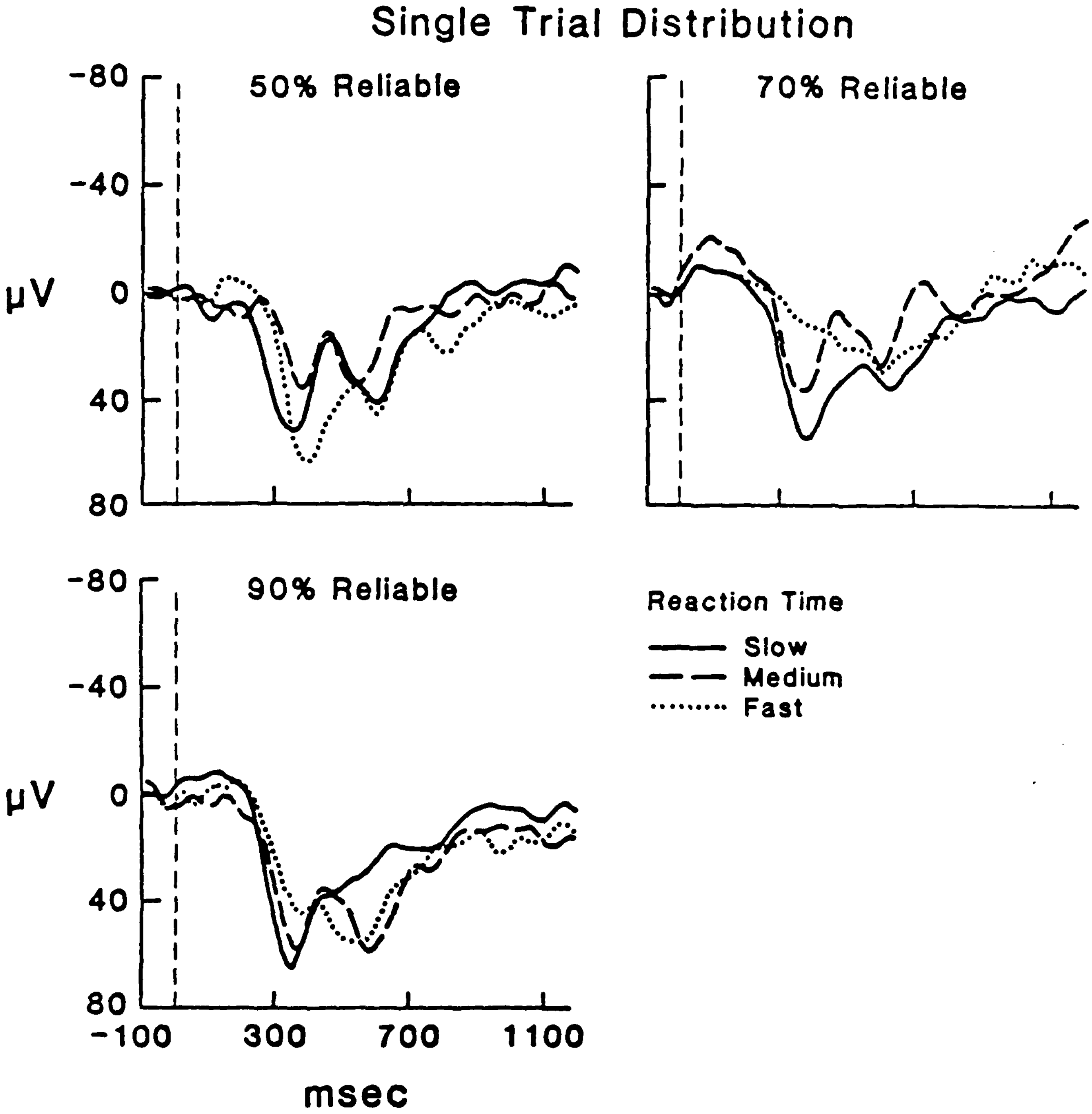

FIGURE 3.5. Single trial ERP distribution of informative stimuli based on RT splits. The RT distributions were divided into 3 parts: fast, medium, and slow. ERPs for each information level were aggregated on the basis of which third of the RT distribution they were associated with.

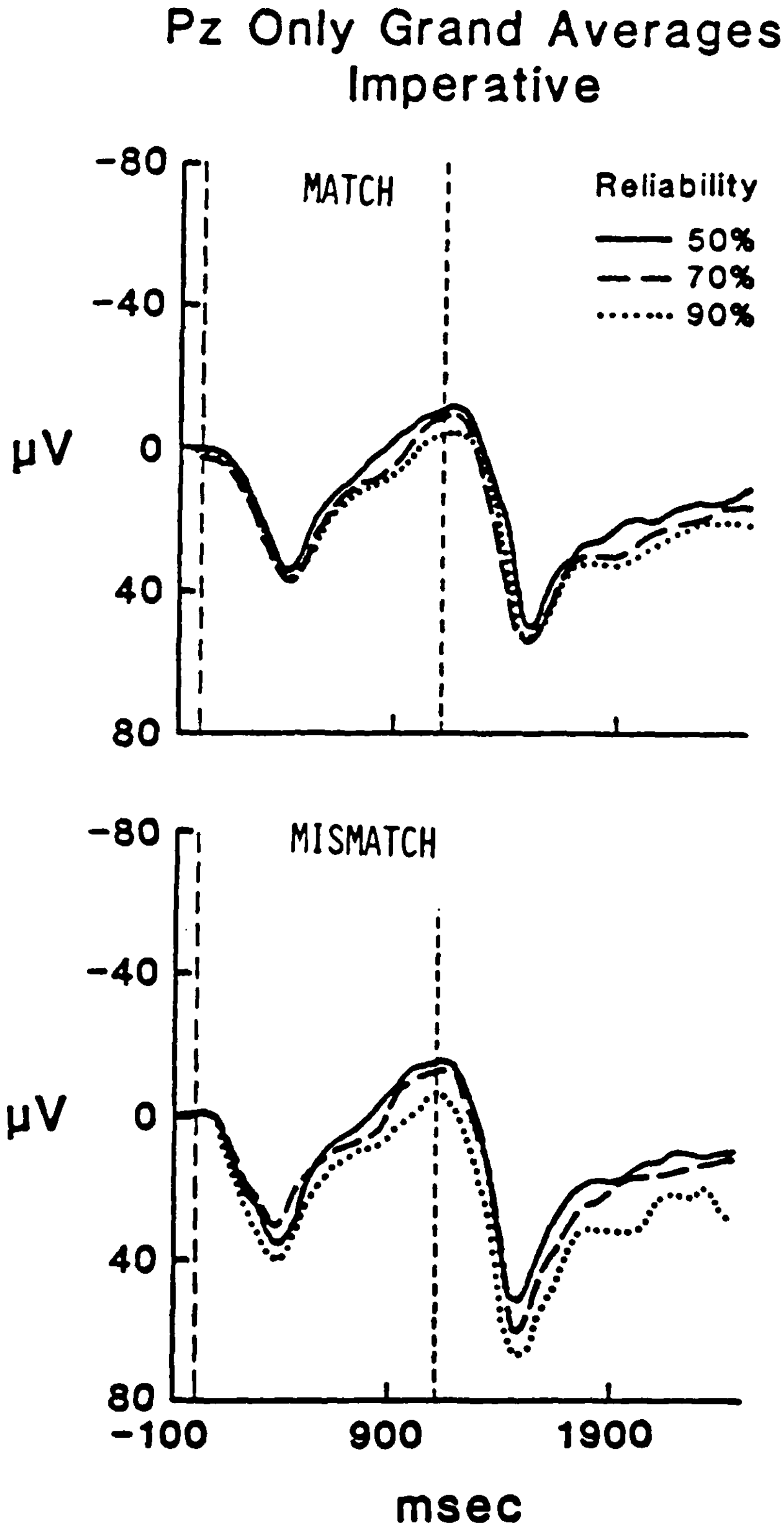

FIGURE 3.6. Grand average for imperative stimuli at Pz for match and mismatch trials. The top plot gives the averages when the imperative stimuli matched the informative stimuli, and the bottom plot shows the averages when the imperative stimuli switched direction from the preceding informative stimuli.

physical identity of the informative stimulus was expected, but the source (and therefore its reliability) was ignored.

Discussion: Experiment 3

The data from this experiment do not conclusively answer the question of whether subjects extract information available to them differentially. The analysis of variance on the mean and median reaction times reveal significant effects for session and for matching. Not surprisingly, subjects' mean RT decreased with practice. Furthermore, the mismatch trials are significantly longer than the match trials. Although the differences among information values are not significant, and the interaction between information value and match condition is not significant, the direction of the change with information value of the match and mismatch RTs is consistent with the hypothesis. RT increased for mismatch trials as the informative value of the stimulus increases, suggesting that subjects were less prepared for a mismatch after a highly reliable cue than after a less reliable cue. Similarly, RT decreases as the informative value of the stimuli increases for the match trials. The question remains as to why the match effect is significant while the interaction for match condition and information value is not. Subjects seem to be extracting expectancy information from the physical form of the cue (up or down arrow) but not differentially, based on cue location.

Since there was no significant effect of information value on RT, it is not surprising that there is a lack of consistency in the P300 data in this study. In fact, the preliminary analysis of individual subject data revealed a great amount of variability. Very few of the subjects exhibited the expected amplitude differences. Furthermore, grand averages did not show a difference in amplitude of the P300 among the levels of information.

Given this display and the scenario that was constructed for the subjects--power plant operators monitoring a subsystem of the total operation--the sequences of events may have seemed too unrealistic for the subjects to place a great degree of confidence in the stability of the system. If the subjects perceived the system as unstable, they may not have been able to build a tenable model of it.

There is one final factor that may have lead to the complete elimination of any effect of information value on P300 amplitude, despite the suggested presence of such an effect in the RT data. To consider the nature of this factor, it is important to reiterate that the amplitude of the P300 elicited by a stimulus may be influenced by three factors: (1)The information value of the stimulus, as demonstrated in Experiments 1 and 2; (2)The absolute probability of the stimulus, and (3)The sequential probability. The latter two effects are typically manifest through the intervening variable of subjective expectancy (Squires, Wickens, Squires and Donchin, 1976; Duncan-Johnson and Donchin, 1977; 1982). While the present experiment was intended to focus attention on the first of these three influences, and the second factor was eliminated by making the six stimuli equiprobable, it is quite possible that the sequential effects on P300 remained. Furthermore the particular constraints built into the transition matrix might have created a situation in which the sequential effects and information value produced opposing and compensating effects on P300 amplitude. The sources of these effects as are follows: Subjects tend to expect stimuli to remain the same, rather than to change (Squires et al., 1976). Thus larger P300's would be predicted by a stimulus that indicated a change of state (reversed arrow direction) than a constant state (repeated arrow direction). However, given the transition probabilities, characterizing the "sluggish" system it is less likely that a change of state will lead to an extreme information value (.9), than to one of lesser value (.7 and especially .5). It follows that most of the state-change stimuli (generating greater P300 amplitudes) will be stimuli of lower information value (predicted by the hypotheses to produce smaller P300s). The present data cannot be easily evaluated to determine the extent of this influence.

GENERAL DISCUSSION

The first research question addressed by the three experiments reported here concerned whether or not people could differentially process information sources of differing levels of information value, and use this information to advantage in updating their expectations regarding the state of the world. Our operational measurement of these expectancies was through reaction time measures: We assumed that the expected state would be rapidly confirmed, while the confirmation of unexpected states would be slow. The answer to this issue ranged from a clear yes in the simplicity of experiment 1, to an apparent no in the complex context of experiment 3, with a "maybe" for the paradigm of intermediate complexity in experiment 2.

In experiment 1, despite its low power, all subjects consistently used the information provided by informative warning stimuli to prepare the appropriate response. This conclusion is not new, and replicates the extensive work on cost-benefit analysis performed by Posner and his colleagues (e.g., Posner, 1978). Paradigm complexity was extended in an important direction in experiment 2, where three, rather than two levels of probabilistic preparation were made available, corresponding to the information values of 0, 4, and 8. In experiment 2, the ability to employ these different levels of information was mixed. Some subjects were clearly able to, and others were not, adopting a strategy more characteristic of the "as if" heuristic. Finally, in experiment 3 it was apparent that very few of the subjects used the graded probabilistic information in an optional fashion, to the extent that no reliable effects of information value on reaction time to the imperative signals was observed.

In this regard the differences in complexity between experiments 2 and 3 are important. Experiment 3 incorporated a more complex series of state transitions that generated the sequences of informative cues. These approximated more realistically the autocorrelated "sluggish" behavior of many real world systems in which the state of the system at one point in time is correlated with its state at another adjacent time point (Jones and Wickens 1986). But the autocorrelation itself may have been confusing or effortful to use, and as noted, may have created sequential effects that negated the information value effects. Secondly, the implicit designation of the cue reliabilities by their spatial location, rather than by numerical value, would have increased the levels of memory. In this light, it is consistent with the premises underlying the "as if" heuristic, that if the costs in mental effort outweigh the benefits of employing differential cue weighting, then the equal reliability "as if" heuristic will be employed.

Recent work in decision making furthermore has begun to be able to objectively document the impact of effort costs on decision strategy selection (e.g., Shugan, 1982; Payne, Bettman & Johnson, 1986). Hence, in interpreting the present results, it may have been that the increased complexity of the information, did not make it worth while to invest effort into its differential use. Subjects found it nearly profitable to prepare equally in response to all three cue values, and were, if anything, influenced in their preparation by the more salient, direct cue of arrow direction.

The second question addressed by the set of experiments concerned the efficacy of the P300 to reflect the differing levels of information extraction. An intrinsic characteristic of the paradigm of course was that our evaluation of the successfulness of the P300 in this endeavor was contingent upon whether or not reaction times validated that the information was in fact being differentially extracted. As noted, this validation was not observed in experiment 3. In experiment 1 on the other hand, the pattern was quite clear. Larger P300's were elicited by informative stimuli than by non-informative ones, despite the fact that the informative ones, as a category occurred twice as often as the non-informative ones, a condition that would actually lead to a smaller P300, if stimulus probability were the only relevant effect. Such a finding supports the view that when information is extracted that bears on the present or future state of the world, an information processing sub-routine manifest in the P300, is invoked. Whether this subroutine engages differentially (and manifests a larger P300) when <u>more</u> information is to be extracted was the second issue examined in Experiment 2. Here the evidence was supportive, although not altogether conclusive. It was indeed the case that larger P300's were observed for the 90% predictive stimuli than for the 70% stimuli, and that the RT data revealed that the preparatory information was extracted from each. The only disconcerting characteristic of the data was the failure of P300 to reliability discriminate between the uninformative .50 stimuli and the low informative .70 stimuli. Here it is possible that low statistical power, was responsible for the lack of effect. Only five subjects data were included in the analysis. An alternative explanation, as noted previously, is that subjects made an initial categorization between high (8) and low (0 and 4) information, reflected in P300, and that the subsequent differential preparation between 0 and 4 took place in an interval following the P300. On the basis of the current data, neither of these possibilities can be rejected.

The final issue addressed by the present research concerns the efficiency of P300 in reflecting differences in information extracted across different regions of a complex visual display. In this regard the current data are are consistent with other positive evidence in the literature (e.g., Wickens Heffley, Kramer and Donchin 1980; Kramer, Wickens and Donchin, 1983), although they do not extend far beyond these previous data. Thus, while experiments 1 and 2 clearly demonstrate effects of information value on P300, both were paradigms in which all information appeared at a single location. Experiment 3 obtained clearly defined P300's to information presented at different locations in the visual field (corresponding to a real world analogy of different physical instruments), but as noted, these P300's were not modulated by information value because the subjects adopted an "as if" strategy. Hence the full richness of P300 as a tool for mapping information extraction across the visual field remains to be established.

References

DeSwart, J. H., Kok, A. & Das-Smaal, E. A. (1981). P300 and uncertainty reduction in a concept-identification task. Psychophysiology, 18, 619-629.

Donchin, E. (1981). Surprise! .... Surprise? Psychophysiology, 18, 493-513.

Donchin, E. & Heffley, E. (1975). Minicomputers in the signal-averaging laboratory. American Psychologist, 30, 299-312.

DuCharme, W. M. (1970). Response bias explanation of conservation human inference. Journal of Experimental Psychology, 85, 66-74.

Duncan-Johnson, C. & Donchin, E. (1977). On quantifying surprise: The variation in event-related potentials with subjective probability. Psychophysiology, 14, 456-467.

Duncan-Johnson, C. & Donchin, E. (1982). The P300 component of the event-related brain potential as an index of information processing. Biological Psychology, 14, 1-52.

Edwards, W. (1968). Conservatism in human information processing. In B. Kleinmuntz (Ed.), Formal representation of human judgment. New York: Wiley, 17-52.

Edwards, W. (1977). *How to use multiattribute utility measurement for social* decision making. IEEE Transactions on Systems, Man & Cybernetics, 17, 326-340.

Edwards, W., Phillips, L. D., Hays, W. L., & Goodman, B. C. (1968). Probabilistic information processing systems: Design and evaluation. IEEE Transactions on Systems, Man and Cybernetics, SSC-4, 248-265.

Einhorn, H. J. & Hogarth, R. M. (1985). A contrast/surprise model for updating beliefs (Technical Report No. 13). Chicago: Center for Decision Research, University of Chicago School of Business.

Fabiani, M., Karis, D. & Donchin, E. (1986). P300 and recall in an incidental memory paradigm. Psychophysiology, 23, 298-308.

Gratton, G., Coles, M.G.H., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. Electroencephalography and Clinical Neurophysiology, 55, 468-484.

Grubin, M. L., Bauer J. A. & Walker, E. C. T. (1976). Alice User's Guide, Natick, Massachusetts.

Harris, R., & Spady, A. A. (1985). Visual scanning behavior. Proceedings of NAECON-85, Dayton, Ohio.

Jasper, H.H. (1958). The ten-twenty system of the International Federation. Electroencephalography and Clinical Neurophysiology, 10, 371-375.

Johnson, E. M., Covenaugh, R. C., Spooner, R. L. & Samet, M. G. (1973). Utilization of reliability measurements in Bayesian inference: Models and human performance. IEEE Transactions on Reliability, 22, 176-183.

Jones, P. & Wickens, C. D. (1986). The display of multivariate information: The effects of auto- and cross-correlation, reliability and heterogeneity. Technical Report No. CPL86-5. Urbana, Illinois: Cognitive Psychophysiology Laboratory, University of Illinois.

Kahneman, D. & Tversky, A. (1973). On the psychology of prediction. Psychological Review, 80, 237-251.

Kramer, A. F. Wickens, C., & Donchin, E. (1985). The processing of stimulus properties : Evidence for dual-task integrality. Journal of Experimental Psychology: Human Perception and Performance, 11, 393-408.

Kramer, A. F., Wickens, C. D. & Donchin, E. (1983). Analysis of processing requirements of a complex perceptual-motor task. Human Factors, 25, 597-621.

Lopes, L. L. (1983). Toward a procedural theory of judgment (Technical Report WHIPP 17). Madison, WI: University of Wisconsin, Department of Psychology, Wisconsin Human Information Processing Program.

Payne, J. W., Bettman, J. R. & Johnson, E. J. (1986). Adaptive strategy selection in decision making (Technical Report 86-1). Durham, N. C.: Duke University School of Business Administration.

Posner, M. I. (1978). Chronometric Explorations of Mind. Hillsdale, N. J.: Erlbaum.

Schum, D. (1975). The weighing of testimony of judicial proceedings from sources having reduced credibility. Human Factors, 1975, 172-203.

Shugan, S. M. (1982). The cost of thinking. Journal of Consumer Research, 7, 99-111.

Slovic, P., Fischoff, B. & Lichtenstein, S, (1977). Behavioral Decision Theory. Annual Review of Psychology. 28, 1-39.

Squires, K. C., Wickens, C. D., Squires, N. K. & Donchin, E. (1976). the effect of stimulus sequence on the waveform of the cortical evoked potential. Science, 193, 1142-1146.

Wickens, E. D., Kramer, A., Vanasse, L. & Donchin, E. (1983). Performance of concurrent tasks: A psychophysiological study of the reciprocity of information processing resources. Science, 221, 1080-1082.

Wickens, C. D. (1984). Engineering Psychology and Human Performance. Columbus, OH: Charles E. Merrill Publishing Company.

Wickens, C. D. (1986). Attention. In P. Hancock (Ed.), Human Factors in Psychology. Amsterdam: North Holland.